Recognition and Enrichment of Archival Documents

D6.1. Binarization and Image Enhancement Tools P1

Ioannis Pratikakis, Konstantinos Zagoris, George Barlas, DUTH
Nikolaos Stamatopoulos, Georgios Louloudis, Basilis Gatos, NCSR
Markus Diem, Stefan Fiel, Florian Kleber, CVL

Distribution:

http://read.transkribus.eu/

READ
H2020 Project 674943

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 674943

Project ref no.	H2020 674943
Project acronym	**READ**
Project full title	**Recognition and Enrichment of Archival Documents**
Instrument	H2020-EINFRA-2015-1
Thematic Priority	EINFRA-9-2015 - e-Infrastructures for virtual research environments (VRE)
Start date / duration	01 January 2016 / 42 Months
Distribution	Public
Contractual date of delivery	31.12.2016
Actual date of delivery	04.12.2016
Date of last update	16.12.2016
Deliverable number	D6.1
Deliverable title	Binarization and Image Enhancement Tools P1
Type	Demonstrator
Status & version	Public & Version 1
Contributing WP(s)	WP6
Responsible beneficiary	DUTH
Other contributors	NCSR, CVL
Internal reviewers	Florian Kleber (CVL), Tobias Grüning (URO)
Author(s)	Ioannis Pratikakis, Konstantinos Zagoris, George Barlas, DUTH Nikolaos Stamatopoulos, Georgios Louloudis, Basilis Gatos, NCSR Markus Diem, Stefan Fiel, Florian Kleber, CVL
EC project officer	Martin Majek
Keywords	Document Image Binarization, Ground Truth, Document Image Enhancement

Table of Contents

Executive Summary

This deliverable reports on the achievements concerning the tasks of document image binarization and document image enhancement at the end of the first year of the READ project that have been realised by three distinct frameworks which correspond to partners NCSR, DUTH and CVL, respectively.

1. NCSR Framework

1.1. Binarization

In the frame of "tranScriptorium" project a novel binarization method was developed which is based on the method of Ntirogiannis *et al*. [Ntirogiannis2014] adapted to handwritten document images. NCSR binarization method has several steps, such as background estimation and image normalization, combination of the global Otsu [Otsu1979] and local Niblack [Niblack 1986] method, contrast calculation, as well as intermediate processing to remove small noisy connected components. The binarization method of Ntirogiannis *et al*. [Ntirogiannis 2014] could miss textual information in an attempt to clear the background from noisy components or bleed-through (Figure 1(a)-(b)). Under those circumstances, some modifications were introduced during the "tranScriptorium" project aiming to preserve textual information (faint characters or low contrast characters) (Figure 1(c)). The main modifications made in [Ntirogiannis 2014] for the development of the NCSR method concern the Otsu threshold selection, the incorporation of a third binarization method ([Gatos2006]) into the combination stage and the noisy connected components removal stage.

There are three different variations of the NCSR binarization method: (i) the default variation which can handle most cases, (ii) the variation which handles bleed-through cases and (iii) the variation which handles the existence of faint characters. The NCSR binarization method is developed in C++ and it is available at github:

https://github.com/Transkribus/NCSR_Tools

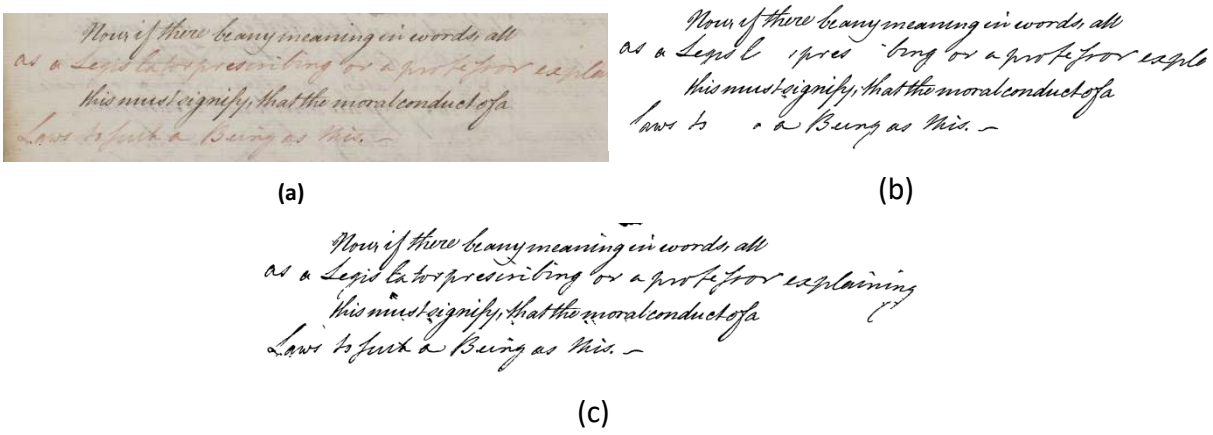

Figure 1: Initial and modified binarization results: (a) original document image, (b) [Ntirogiannis 2014] binarization method, (c) NCSR binarization method.

2. DUTH Framework

2.1. Binarization

In the frame of "READ" project a binarization method was adopted from the "tranScriptorium" project (See Section 1.1). The modifications include the merge of the three different basic variations (Default, Bleed-through and Faint characters') into a single generic mode along with further optimizations with respect to the selection of various thresholds. The binarization method is developed in C++11 and is available at github under LGPL-3.0:

https://github.com/Transkribus/VCG-DUTH-ImageBinarizationTool

Figure 2(b) shows an example of the binarization algorithm.

(a) (b)

(c) (d)

Figure 2: (a) Original document Image, (b) Binarization result, (c) Image Enhancement, (d) Binarization result of the enhanced image shown at (c).

2.2. Image Enhancement

In the frame of "READ" project, document image enhancement tools have been implemented to be used as a preprocessing stage for the binarization algorithm. After experimental work, we have built the enhancement tool which comprises two distinct modules.

The first module is an anisotropic diffusion process or also called Perona–Malik diffusion [Perona1987]. It is a procedure aiming at reducing image noise without removing significant parts of the image content - typically edges.

The second module is inspired by adopting the characteristics of the ganglion cells of the Human Visual System (HVS) [Nelson2004, Vonikakis2011]. It can deal with various types of degradations, such as uneven illumination, shadows, low contrast, smears, and heavy noise densities.

The document image enhancement tool is developed in C++11 and it is available at github under LGPL-3.0:

https://github.com/Transkribus/VCG-DUTH-Enhancement

Figure 2 (c) – (d) shows an example of the enhancement process and its impact on the corresponding binarization in qualitative terms.

3. CVL Framework

The developed *CVL Framework* contains basic image processing algorithms and methodologies for Document Image Analysis. It is developed within the project and is the basis for the WP of CVL. It contains a binarization and a skew estimation (document image preprocessing). The framework is developed in C++ and available at github under LGPL-3.0:

https://github.com/TUWien/ReadFramework

https://github.com/TUWien/ReadModules

ReadModules contain C++ plugins to execute the binarization and skew estimation. The following sections summarize the binarization and skew estimation.

3.1. Binarization

The developed binarization is based on Su et al. [SU2010] which uses the local minimum and maximum to provide a normalized contrast image to suppress the background variation. The proposed method in Su et al. [SU2010] is extended by an estimation of the page area within the image. Additionally, a weighting of the foreground is possible to suppress noise in document images (similar to the method proposed in Kleber et al. [Kleber2011]).

Figure 3 shows a screenshot of the binarised image in the READ framework and the original image of the H-DIBCO 2016 dataset.

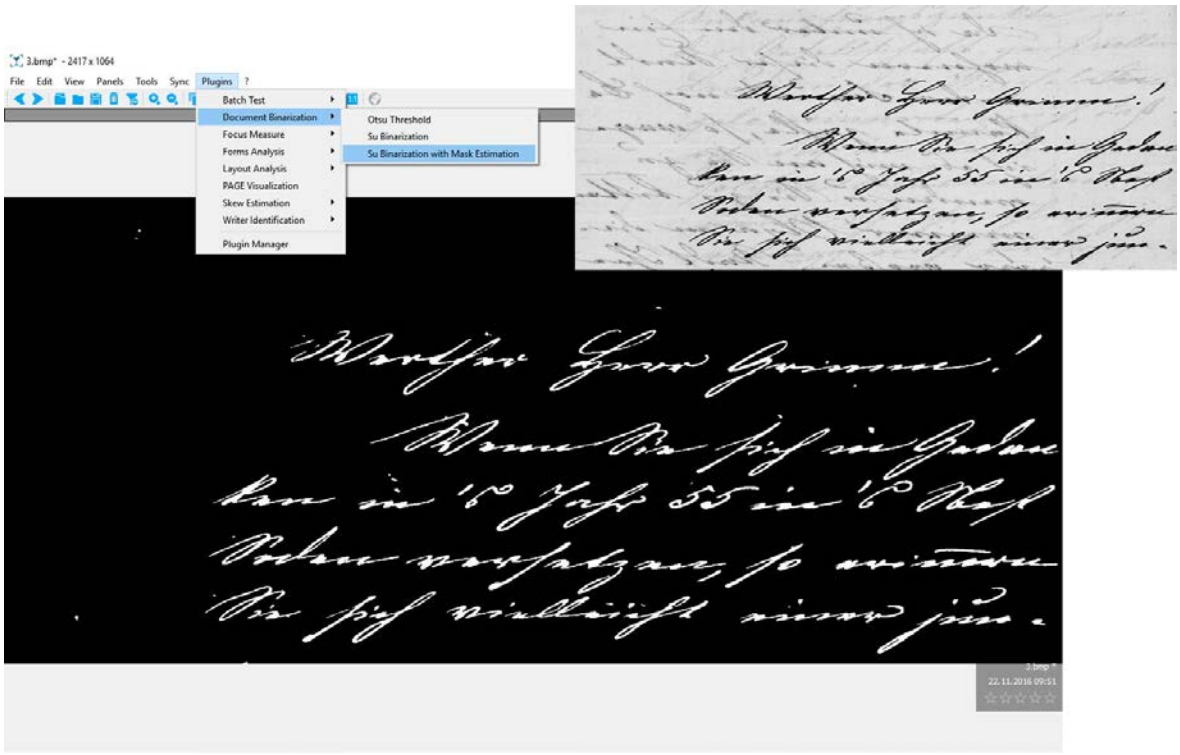

Figure 3: Original image and binarised version of an Image of H-DIBCO 2016 using the binarization from CVL framework.

3.2. Skew Estimation

The skew estimation is based on Koo and Cho [Koo2013]. The methodology uses a salient line detector for natural images. For the READ framework the skew estimation of [Koo2013] is evaluated and the parameters are adapted according the need of document images. Figure 4 shows a screenshot of the skew estimation's result applied to an image of the Document Image Skew Estimation Contest (DISEC) [Papandreou2013], which was rotated by 9.13°.

The method was evaluated on the entire DISEC dataset using the DISEC metric. The following table shows the results:

AED	**0.395°**
CE	0.392
Top80	0.11°
Median Error	0.13°

In contrast to other state-of-the-art skew estimation methodologies, the proposed algorithm can correctely deskew documents which are rotated up to 180°. As future work, the current method will be combined with a gradient based approach (similar to [Diem2012]).

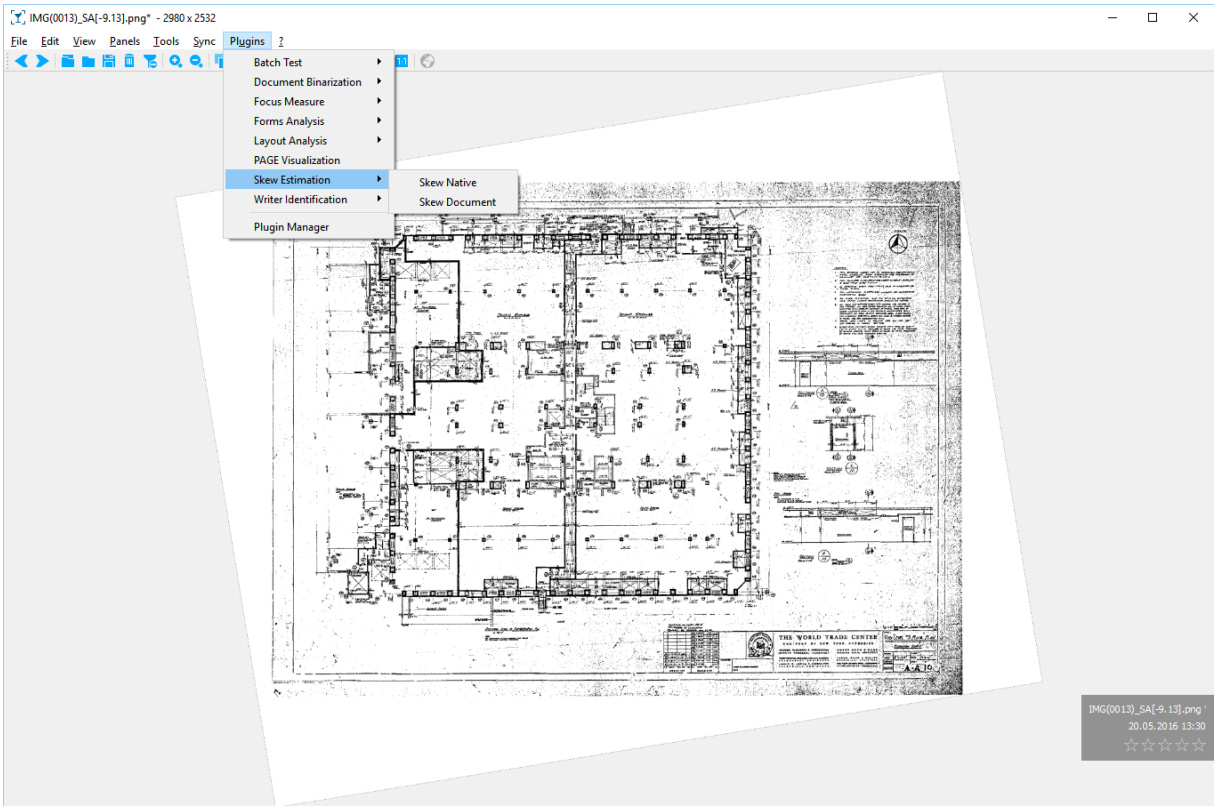

Figure 4: Skew Estimation of the READ framework applied to an image of the DISEC [Papandreous2013] (rotated by 9,13°).

4. Evaluation - Conclusive remarks

The binarization frameworks described in the aforementioned Sections have been evaluated on the latest H-DIBCO2016 [Pratikakis2016] dataset. Table 1 shows the performance achieved from each method tested for each evaluation metric of the H-DIBCO 2016 benchmark. Detailed description of the metrics is given in [Pratikakis2016].

As the experimental results indicate, method (ii) of the NCSR binarization framework, which handles bleed-through cases, outperforms the other two variations, namely method (i), (iii), respectively. In addition, it is ranked first among all methods in all frameworks in terms of FM and PSNR evaluation measures.

Concerning DRD evaluation measure, the best performance has been achieved by the binarization of the DUTH framework. This method has been also tested by using the enhancement operation described in Section 2.2 as a preprocessing step. As shown in Table 1, this resulted in achieving the best performance among all methods in all frameworks in terms of the pseudo-FM metric which has been introduced in [Ntirogiannis2013] to compensate for the subjectivity of the ground truth. The improved performance achieved with the use of the proposed enhancement sheds light on the added value of using such a preprocessing step and quantifies the improved results of the enhancement stage that has already been shown for example cases, as in Fig. 2c, in qualitative terms.

Table 1: Performance evaluation of document image binarization for the H-DIBCO2016 dataset.

Framework	*FM*	*pseudo-FM*	*PSNR*	*DRD*
DUTH	90.62	90.95	19.04	**3.86**
DUTH (/w enhancement)	86.64	**91.39**	17.67	4.67
NCSR – method (i)	88.67	87.61	18.42	5.80
NCSR – method (ii)	**90.84**	89.79	**19.31**	4.00
NCSR – method (iii)	84.40	83.49	16.52	9.29
CVL	*85.89*	*90.72*	*17.64*	*5,43*

5. References

[Diem2012] Diem, Markus, Florian Kleber, and Robert Sablatnig. 2012. "Skew Estimation of Sparsely Inscribed Document Fragments." Proceedings of the International Workshop on Document Analysis Systems (DAS).

[Gatos 2006] B. Gatos, I. Pratikakis, and S.J. Perantonis, "Adaptive degraded document image binarization," Pattern Recognition, vol. 39, no. 3, pp. 317–327, Mar. 2006.

[Klebel2011] Kleber, Florian, Markus Diem, and Robert Sablatnig. 2011. "Scale Space Binarization Using Edge Information Weighted by a Foreground Estimation." Proceedings of the 11th International Conference on Document Analysis and Reconstruction (ICDAR 2011). 854-858.

[Koo2013] Koo, Hyung Il, and Nam Ik Cho. 2013. "Skew estimation of natural images based on a salient line detector." Journal of Electronic Imaging 22(1) 013020-1-6.

[Nelson2004] Nelson R, Kolb H (2004) ON and OFF pathways in the vertebrate retina and visual system. In: Chalupa LM, Werner JS (eds) The visual neurosciences. MIT Press, Cambridge, pp 260–278

[Niblack 1986] W. Niblack, An Introduction to Digital Image Processing. Englewood Cliffs, NJ: Prentice Hall, 1986, pp. 115–116.

[Ntirogiannis2013] K. Ntirogiannis, B. Gatos and I. Pratikakis, "Performance Evaluation Methodology for Historical Document Image Binarization," in *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 595-609, Feb. 2013.

[Ntirogiannis 2014] K. Ntirogiannis, B. Gatos and I. Pratikakis, "A Combined Approach for the Binarization of Handwritten Document Images", *Pattern Recognition Letters - Special Issue on Frontiers in Handwriting Processing*, vol. 35, no.1, pp. 3-15, Jan. 2014.

[Otsu 1979] N. Otsu, “A Thresholding Selection Method from Gray-level Histogram,” IEEE Trans. Syst., Man, Cybern., vol. 9, no. 1, pp. 62–66, 1979.

[Papandreou2013] Papandreou, A., Basilis Gatos, George Louloudis, and N. Stamatopoulos. 2013. “ICDAR2013 Document Image Skew Estimation Contest (DISEC’13).” 12th International Conference on Document Analysis and Recognition. 1444-1448.

[Perona1987] Pietro Perona and Jitendra Malik. 1987. "Scale-space and edge detection using anisotropic diffusion". Proceedings of IEEE Computer Society Workshop on Computer Vision,. pp. 16–22.

[Pratikakis2016] Pratikakis, Ioannis, Konstantinos Zagoris, George Barlas, and Basilis Gatos. 2016. “ICFHR 2016 Handwritten Document Image Binarization Contest (H-DIBCO 2016).” 15th International Conference on Frontiers in Handwriting Recognition. Shenzen. 620-623.

[Su2010] Su, Bolan, Shijian Lu, and Chew Lim Tan. 2010. “Binarization of Historical Document Images Using the Local Maximum and Minimum.” 9th IAPR International Workshop on Document Analysis Systems. Boston. 159-166.

[Vonikakis2011] Vonikakis, V., Andreadis, I., & Papamarkos, N. (2011). Robust document binarization with OFF center-surround cells. Pattern Analysis and Applications, 14(3), 219-234.